Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 9
AIDML 2025

Progress of Pathological Image Analysis Based on deep Learning

Zibing Hua *

Facility of Software Engineering, Southeast University, Nanjing, 211189, China

* Corresponding Author Email: 213223620@seu.edu.cn

Abstract. Pathological image analysis plays a crucial role in the diagnosis and treatment of diseases. However, due to its complexity and diversity, traditional methods are difficult to meet the growing demand. A novel method for analyzing pathological images has emerged in recent years because of advancements in deep learning technology, particularly the usage of convolutional neural networks (CNNs). This study compared the performance of various models through a systematic literature review and investigated the application status of deep learning in pathological image analysis, including pathological image classification, object detection and segmentation, and image generation and enhancement. According to studies, deep learning techniques greatly increase the precision and effectiveness of pathological picture analysis. This work not only emphasizes the significant potential of deep learning technology in enhancing the efficiency of pathological image processing but also brings out its practical application value in clinical practice. Furthermore, to further optimize the integration of deep learning into workflows for pathological image analysis, this study also analyzes the remaining obstacles and future objectives.

Keywords: Deep learning, pathological image analysis, convolutional neural network.

1. Introduction

Analysis of pathological images is essential for both diagnosing illnesses and developing treatment plans. Pathological image analysis has witnessed tremendous advancements in recent years due to the emergence of deep learning technologies, particularly the widespread usage of CNNs. Echle et al. pointed out that deep learning has a promising application as a new generation of clinical biomarkers in cancer pathology [1]. Suganyadevi et al. highlighted the great potential of deep learning in medical image analysis, including but not limited to cancer detection, tissue classification, etc. [2]. Despite the remarkable progress, Van der Laak et al. also pointed out the challenges in the process of moving from laboratory to clinical application, such as model interpretability and data requirements [3]. There are still some research limitations in the existing literature, such as how to effectively apply deep learning to the actual pathology workflow and how to improve the transparency and interpretability of the model.

This paper focuses on the research related to pathological image analysis based on deep learning, aiming to explore the latest application and development trend of deep learning technology in pathological image analysis by comprehensively analyzing the relevant literature. This paper introduces several major deep learning methods and their specific implementation details in pathological image processing and discusses the effectiveness and limitations of these methods. Finally, it can provide a valuable reference for future research and promote the further development of this field.

2. Pathological Image Analysis Features

Pathological image analysis is an important field of medical image analysis that has the following characteristics. First, image complexity is high; the pathology image (such as whole slice images, WSI) usually has a high resolution (up to billions of pixels) and contains abundant details. The second image also has the characteristics of diversity, images may come from different dyeing methods (such as H&E staining, immunohistochemical, fluorescent staining), each of the staining methods to present information is different, the form, arrangement, and distribution of tissues and cells with high complexity, such as the nucleus, glands, blood vessels, such as the forms of the structure. In addition, pathological images usually have a large scale of data, which usually needs to be processed in blocks. The above characteristics result in the low applicability of traditional methods. Deep learning models, especially CNN, can learn and extract features better and reduce the dependence on large-scale labeled data. The key features and parameters of pathological image recognition are shown in Table1.

Table 1. Key features and parameters of pathological image recognition.

Types of pathological images	Key features	Identifying parameters	References/Remarks
Photographic images of gross specimens	Color and morphological features	Color distribution and morphology	Kather et al. [4]
Microscopic image of tumor cells	Color statistical features	Histogram statistics	Madabhushi et al. [5]
	Texture features	Wavelet packet decomposition energy, fractal dimension	Sulliva et al. [6]
	Shape feature	7-dimensional moment invariants	Litjens et al. [7]
Tissue slice images	Grayscale statistical features	Gray distribution parameters	Cruz-Roa et al. [8]
Image of cells after staining treatment	Texture complexity	Fractal dimension	Veta et al. [9]
	Nucleo-plasmic ratio characteristics	The area ratio of the nucleus to cytoplasm	Gurcan et al. [10]

3. Research on Deep Learning Technology in Pathological Image Analysis

Deep learning provides a rich variety of solutions for the analysis process of pathological images. By integrating deep learning technology into traditional analysis methods (such as color and texture-based feature extraction and classification technology [11], and morphology-based image segmentation technology [12]), it can improve the generalization degree of complex and diverse tissue images and optimize the efficiency of the algorithm. Researchers in the field of pathological image analysis using deep learning are currently primarily separated into the following various elements.

3.1. Application of CNN in Pathological Image Classification:

The goal of pathological image classification is to automatically identify the lesion type and grade or predict the prognosis of patients from tissue sections to provide support for clinical diagnosis and treatment. The following are the main research methods of deep learning technology in this field.

3.1.1. Application of CNN in Pathological Image Classification:

In 2022, Li et al. proposed a Multi-Task Learning (MTL) framework for pathological image classification. Through multi-task learning (MTL), the underlying convolutional layer of the model shares weights between different tasks (e.g., classification, image verification), reduces redundant parameters and enhances feature generalization ability. At the same time, algorithms such as GradNorm are used to adaptively adjust the loss weights of different tasks to avoid optimization bias caused by differences in task difficulty.

By using the above techniques, the image classification effect has been significantly improved. The contrastive learning strategy is used to narrow the feature distance between samples from the same class and expand the difference between samples from different classes. For example, in breast cancer classification, this strategy reduces the intra-class variance of the same type of malignant cells by 18% [13]. In addition, jointly optimizing multiple related tasks (e.g., classification + localization) prevents the model from overfitting the noise of a single task and improves generalization. According to experiments, the multi-task model's test set overfitting rate is 7.3% lower than the single-task models [13].

On the breast cancer data set (e.g., BreakHis), multitasking CNN's overall classification accuracy of 96.2%, a single task ResNet - 50 increases by 4.8%. The AUC value of malignant and benign classification is improved from 0.91 to 0.97, indicating that the model is more discriminative in distinguishing borderline cases (such as atypical hyperplasia). For cross-dataset generalization ability, transfer experiments on Camelyon16 (lymph node metastasis detection) and TCGA (multiple cancer classification) datasets show that the multi-task model's average F1 score is 0.89, 0.15 greater than the single-task models [14].

3.1.2. Application of transfer learning in pathological image classification:

However, CNN relies on a large amount of labeled data to fully train the deep network. However, the cost of pathological image annotation is high, and the samples of rare diseases are scarce (e.g., only hundreds of cases in the Camelyon16 dataset), which makes the model easy to overfit. Transfer learning can reduce the sample size required for pathological image training by reusing pre-trained features. For example, in breast cancer classification, 83.5% accuracy can be achieved by fine-tuning only 10% of the layers [11].

Transfer learning adjusts the model by adjusting the parameters of the convolutional layer and the parameters of the fully connected layer. Specifically, the shallow convolutional layers are frozen, and the deep convolutional layers are fine-tuned to adapt to pathological specific features. Moreover, the pre-trained model's classification head is changed (e.g., the 1000 class output of ImageNet is changed to the binary classification of benign/malignant breast cancer), and the weights of the fully connected layer are retrained.

Experiments show that the classification effect is improved under the transfer learning method, which is reflected in the following aspects: First, the accuracy is improved. The multi-level transfer learning based on DenseNet reaches 83.5% on the BreakHis breast cancer dataset, which is 12% higher than that of traditional CNN (such as DenseNet without transfer) [11]. The second is the enhancement of cross-domain ability. In the cross-domain task from Camelyon16 (lymph node metastasis) to TCGA (multiple cancer types), the average F1 score of transfer learning reaches 0.89, this surpasses the single-domain training models by 0.15. The last advantage is the small sample scene; in only 500 copies marked image of rare disease classification task, migration study the AUC value is $0.92, significantly higher than that of traditional CNN 0.78 [15].

3.1.3. Application of transfer learning in pathological image classification:

The core advantages of multi-label classification can significantly improve the logic of classification by modeling label sequence dependencies by RNN or attention mechanisms (e.g., CNN-RNN framework) or using graph models to describe label co-occurrence probabilities. - for example, CNN RNN framework in predicting "tissue adjacent to carcinoma", through the guidance of RNN hidden states model focused on "fibrosis associated" label, F1 scores increase 9%.

In addition, through the dynamic attention mechanism (such as implicitly adjusting the attention region), the model focuses on different regions when predicting different labels. For example, in "vascular invasion", the focus module automatically strengthens the blood vessel edge character, making sensitivity increase by 15%. In addition, semantic redundancy is used to improve generalization, and classification parameters of semantically similar labels are shared through Joint Embedding space, which reduces the computational cost and improves generalization ability. Experiments reveal that the average accuracy of the proposed strategy is enhanced by 12% in cross-dataset transfer (e.g., NUS-WIDE to TCGA)

3.2. Application in Object Detection and Segmentation

Deep learning techniques, especially object detection and segmentation methods, are playing an increasingly important role in pathological image analysis. Object detection aims to identify and locate specific structures or lesion regions in tissue slices, and commonly used methods include Mask R-CNN-based algorithms. These methods automatically identify objects of interest and accurately delineate their boundaries by training deep convolutional neural networks. Image segmentation goes a step further and requires that each pixel in the image be classified to distinguish between different tissue types or lesion states. Using the improved Mask R - CNN model is precise and efficient for completing tasks for detecting and segmenting the nucleus. The progress of these technologies has greatly improved the speed and accuracy of pathological diagnosis, which provides strong support for precision medicine.

3.2.1. Application of region-based convolutional neural network (R-CNN) in object detection and segmentation:

R-CNN generates the candidate region (region proposals), and then for convolution operation to extract the features of each area, finally returns classifying and bounding box. Improved versions, such as Faster R-CNN and YOLO, further improve detection speed and accuracy.

Faster R-CNN introduces an anchor mechanism, presets reference frames of different sizes and proportions, and dynamically generates high-quality candidate regions through convolutional feature maps, which significantly reduces redundant calculations. FPN (Feature pyramid network) or cross-layer connection is used to fuse feature maps at different levels to improve the detection ability of small targets (such as cancer cell nuclei). Improved Faster - R - CNN (a combination of attention mechanism) on Camelyon16 dataset mAP up to 92.3%, up 8.5% than the traditional methods [15].

3.2.2. Application of U-Net in object detection and segmentation:

The core advantage of U-Net over R-CNN is that its symmetric encoder-decoder structure and skip connection directly retain multi-scale features and do not need to rely on region proposal Generation (RPN), which is more suitable for the segmentation of dense small objects (such as nuclei) [16].

The encoder depth and the initial number of channels are the primary influencing factors for U-Net. Specifically, increasing the number of layers (e.g., ResNet-34 as the encoder) can improve the feature abstraction ability. Expanding the number of channels (e.g., 64→128) enhances sensitivity to small objects but increases the computational cost. In 2020, Ozan Oktay et al. introduced Attention U-Net [10], which proposes an attention mechanism based on U-Net to enable the network to focus on the target location [16]. In addition, there are also TransUNet proposed by Chen et al., and multi-ResUNet proposed by Alom et al., which optimize traditional methods in pathological image segmentation tasks [17].

The improved U-Net (combined with Transformer) achieves a Dice coefficient of 0.92 and a classification accuracy of 98.1% on the MoNuSeg dataset, which is 9% higher than that of Mask R-CNN. The Dice coefficient of Attention U-Net on the Camelyon16 dataset is 0.89, and the false positive rate is 34% lower than that of Faster R-CNN [2].

3.2.3. Application of Mask R-CNN in object detection and segmentation:

Using a fully convolutional network, Mask R-CNN directly generates a target mask and adds a segmentation branch to Faster R-CNN to enable end-to-end instance segmentation (e.g., accurate contour extraction of pathological objects such as nuclei and glands) [14]. Another advantage of Mask R-CNN is that it replaces ROI Pooling of R-CNN, eliminates quantization error, preserves spatial details of targets, and significantly improves the segmentation accuracy of small targets (such as tumor cells). In 2020, an automated system based on Mask R-CNN was developed for lymphocyte segmentation in whole-slide imaging. Studies have shown that the proposed system can efficiently and accurately identify lymphocytes in large-scale pathological images, which is helpful in accelerating the research process of cancer immunotherapy.

3.3. Application in Image Generation and Enhancement

Traditional medical image processing techniques are being revolutionized in the field of pathological image analysis by deep learning-based image generation and enhancement techniques. Image generation techniques, such as Generative adversarial networks (GANs), can generate high-quality pathological images based on a small number of samples, greatly enrich the training dataset, and help to improve the generalization ability of the model. Such technological advances are of great significance for realizing personalized medicine, promoting pathological research, and improving patient treatment outcomes.

3.3.1. Application of generative adversarial networks (GAN) in image generation and enhancement:

GANs have shown great potential in the generation and enhancement of pathological images. A discriminator and a generator are the two primary parts of gans; the discriminator is in charge of differentiating between created and actual images, while the generator is entrusted with producing realistic images from random noise or low-quality input. Both are jointly optimized through adversarial training until the generator can produce high-quality images that are difficult to distinguish.

By adjusting the adversarial strength of the generator and the discriminator (such as the gradient penalty coefficient λ in Wasserstein GAN). The Perceptual Loss or L1/L2 loss is introduced to constrain the semantic consistency between the generated image and the real image.

In pathology, GAN can be used to generate new and synthetic pathological images to compensate for the problem of insufficient datasets. For example, models such as PathologyGAN can generate high-resolution tissue slice images, which can help expand the training dataset and improve the performance of deep learning models.

3.3.2. Application of GAN in image generation and enhancement:

Image super-resolution technology converts low-resolution images into high-resolution images through deep learning models (e.g., SRCNN, ESRGAN). In pathology, this helps to improve image details and facilitate more accurate analysis. Gans can be used to enhance the quality of existing pathology images. For example, image resolution can be improved through Super-resolution GAN (SRGAN) technology, making cellular structures more clearly visible [18]. CycleGAN can be used to convert between images of different modalities to help improve image contrast and details to assist doctors in making a more accurate diagnosis.

4. Challenges and Development Directions of Deep Learning for Pathological Image Analysis

4.1. Challenges

(1) Hardware requirements caused by data quality: Various sources of pathology images lead to inconsistent data formats, resolutions, and color standards, and huge storage requirements of pathological sections, which require high hardware infrastructure.

(2) Low model interpretability: The lack of transparency in the decision-making process of deep learning makes it difficult for doctors to understand the reasoning logic of AI, especially when complex diagnoses (e.g., tumor classification) are involved, which may affect clinical adoption.

(3) Insufficient cross-domain generalization ability: The model performs well on a single dataset (e.g., images collected by a specific hospital or device), but its performance significantly decreases in data that cross institutions, diseases, or staining methods.

4.2. Development

(1) Data sharing and federated learning: Improve the generalization ability of the model by realizing cross-institutional data collaboration through private computing technologies (e.g., secure multi-party computation).

(2) Visual reasoning and decision support: Develop an interpretable framework (e.g., attention mechanism, heatmap annotation) to visualize the AI lesion recognition process.

(3) Cloud services and edge computing: Using lightweight models (e.g., knowledge distillation technology) to reduce the demand for local computing power and promote the implementation of AI in primary hospitals.

5. Conclusion

This study explored the application status and challenges of deep learning in pathological image analysis through a systematic literature review. Research has demonstrated that deep learning techniques greatly increase the precision and effectiveness of pathological picture analysis. In particular, the application of CNNs has shown excellent performance in many aspects, such as cancer detection and tissue classification. In addition, advances in GANs and super-resolution technology provide new solutions for the generation and enhancement of pathological images.

This study not only highlights the great potential of deep learning technology to improve the efficiency of pathological image analysis but also points out its practical application value in clinical practice. For example, the successful application of multi-task learning frameworks and transfer learning strategies has enabled the training of efficient models even when there is a lack of information. Meanwhile, the attention mechanism and the improved U-Net architecture further improve the accuracy of image segmentation tasks. The development of these technologies is of great significance for improving the medical diagnosis process, improving the early detection rate of diseases, and formulating personalized treatment plans.

Current issues focus on the insufficient standardization of pathological imaging data, the low interpretability of the model, and the weak cross-domain generalization ability. The development direction emphasizes promoting the clinical credibility of pathological AI and the popularization of primary medical care through a data collaboration framework, interpretable technologies (e.g., attention mechanism), and lightweight model deployment (e.g., federated learning, cloud services).

References

[1] Echle, A., Rindtorff, N. T., Brinker, T. J., et al.: 'Deep learning in cancer pathology: A new generation of clinical biomarkers', British Journal of Cancer, 2021, 124(4), pp. 686-696.

[2] Suganyadevi, S., Seethalakshmi, V., Balasamy, K.: 'A review on deep learning in medical image analysis', International Journal of Multimedia Information Retrieval, 2022, 11(1), pp. 19-38.

[3] Van der Laak, J., Litjens, G., Ciompi, F.: 'Deep learning in histopathology: the path to the clinic', Nature Medicine, 2021, 27(5), pp. 775-784.

[4] Kather, J. N., et al.: 'Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer', Nature Medicine, 2019, 25(7), 1054-1056.

[5] Madabhushi, A., Lee, G.: 'Image analysis and machine learning in digital pathology: Challenges and opportunities', Medical Image Analysis, 2016, 33, pp. 170-175.

[6] Sullivan, A.C., Hunt, J.P., Oldenburg, A.L.: 'Fractal analysis for classification of breast carcinoma in optical coherence tomography', Journal of Biomedical Optics, 2011, 16(6), pp. 066010.

[7] Litjens, G., et al.: 'A survey on deep learning in medical image analysis', Medical Image Analysis, 2017, 42, pp. 60-88.

[8] Cruz-Roa, A., et al.: 'Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks', Proc. SPIE Medical Imaging, 2014, 9041, 904103.

[9] Veta, M., Pluim, J.P., van Diest, P.J., Viergever, M.A.: 'Breast cancer histopathology image analysis: a review', IEEE Trans Biomed Eng, 2014, 61(5), pp. 1400-1411.

[10] Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., Yener, B.: 'Histopathological image analysis: a review', IEEE Rev Biomed Eng, 2009, 2, pp. 147-171.

[11] Wang, W. T., Zheng, Y.: 'Breast cancer image classification based on deep transfer learning', Data Mining, 2022, 12(2), pp. 192-202.

[12] Haralick, R. M., Shanmugam, K., Dinstein, I.: 'Textural Features for Image Classification', IEEE Transactions on Systems, Man, and Cybernetics, 1973, 3(6), pp. 610-621.

[13] Li, L., Pan, X., Yang, H., et al.: 'Multi-task deep learning for fine-grained classification and grading in breast cancer histopathological images', Multimedia Tools and Applications, 2020, 79, pp. 14509–14528.

[14] Oktay, O., et al.: 'Attention U-Net: Learning Where to Look for the Pancreas', Medical Image Analysis, 2020, 62, 101663.

[15] Li, Y., Song, P.: 'Research progress of transfer learning in medical image classification', Journal of Image and Graphics, 2012, 27(3), pp. 672-686.

[16] Xia, F., et al.: 'Graph Learning: A Survey', IEEE Transactions on Artificial Intelligence, 2021, 2(2), pp. 109-127.

[17] Chen, J., et al.: 'TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation', arXiv preprint arXiv:2102.04306, 2021.

[18] Wang, Y., et al.: 'Enhanced pathology image quality using GAN-based super-resolution techniques', Journal of Digital Imaging, 2021, 34(5), pp. 1076-1087.